

J Math Imaging Vis (2013) 47:35–47
DOI 10.1007/s10851-013-0416-9

Statistical Model of Shape Moments with Active Contour Evolution for Shape Detection and Segmentation

Yan Zhang · Bogdan J. Matuszewski · Aymeric Histace · Frédéric Precioso

Published online: 26 January 2013
© Springer Science+Business Media New York 2013

Abstract This paper describes a novel method for shape representation and robust image segmentation. The proposed method combines two well known methodologies, namely, statistical shape models and active contours implemented in level set framework. The shape detection is achieved by maximizing a posterior function that consists of a prior shape probability model and image likelihood function conditioned on shapes. The statistical shape model is built as a result of a learning process based on nonparametric probability estimation in a PCA reduced feature space formed by the Legendre moments of training silhouette images. A greedy strategy is applied to optimize the proposed cost function by iteratively evolving an implicit active contour in the image space and subsequent constrained optimization of the evolved shape in the reduced shape feature space. Experimental results presented in the paper demonstrate that the proposed method, contrary to many other active contour segmentation methods, is highly resilient to se-

vere random and structural noise that could be present in the data.

Keywords Active contour · Statistical shape model · Segmentation · Shape detection

1 Introduction

Originally proposed in [13], active contour models have achieved enormous success in image segmentation. The basic idea of active contour is to iteratively evolve an initial curve towards the boundaries of target objects. Classical curve evolution is normally driven by a combination of internal forces, determined by the geometry of the evolving curve, and external forces induced from an image. A segmentation method using active contour is usually based on minimizing a functional defined in such a way that for curve close to the target object boundary it has small value.

Introduction of a prior shape constraint into the image segmentation functional has recently become the focus of intensive research [6, 7, 12, 14, 16, 19, 24]. The early work on this problem has been done by Cootes et al. [4]. Their method is based on principal component analysis (PCA) calculated for landmarks selected for a training set of shapes which are assumed to be representatives of the shape variations. The method is implemented in the parametric active contour framework, with results strongly depending on the quality of the selected landmarks.

Leventon et al. [17] considered introduction of prior shape information using level set based representation, where landmarks are replaced by signed distance functions calculated for the contours in the training data set, providing hence an intrinsic and parametrization free shape

This work was supported by the Engineering and Physical Sciences Research Council [Grant number EP/D077540/1].

Y. Zhang · B.J. Matuszewski (✉)
School of Computing, Engineering and Physical Sciences,
University of Central Lancashire, Preston PR1 2HE, UK
e-mail: bmatuszewski1@uclan.ac.uk

Y. Zhang
e-mail: zhangyan_79@hotmail.com

A. Histace
ETIS-UMR8051 CNRS, ENSEA, University Cergy-Pontoise,
95014 Cergy-Pontoise, France
e-mail: aymeric.histace@ensea.fr

F. Precioso
Laboratoire I3S, UMR7271-UNS CNRS, Université Nice Sophia
Antipolis, 06900 Sophia Antipolis, France
e-mail: precioso@polytech.unice.fr

model. However, it was demonstrated that, linear combinations of signed distance functions do not necessarily result in a signed distance function, and therefore possibly compromise the quality of the solution. Furthermore, all these methods effectively assume that the shape prior has a Gaussian distribution. As a result, these methods cannot handle multi-modal shape distributions and thus are restricted to the segmentation of target objects with limited shape variabilities.

Instead of using evolution of active contour to search optimum in the image space, Tsai et al. [25] proposed a method to directly search solution in the shape space which is built by the signed distance functions of aligned training images and reduced by PCA. In their paper, a few cost functions are proposed and their derivatives with respect to eigen-shape weights and to pose parameters are given, so that the steepest descent algorithm can be applied. In [10], Fussenegger et al. apply a robust and incremental PCA algorithm on binary training masks of the object(s) to define an active shape model which is then “embedded” in a level set implementation. Segmentation (or tracking) is computed using pre-trained shape model, then PCA representation is updated using this result in order to improve next iteration of segmentation process. Although this self-improving “looping process” between the image space and the shape space is interesting, PCA of binary training masks requires that these training examples are aligned before learning the implicit shape model. The major limitation of all these methods is the implicit assumption of uniform distribution in the shape space.

Recently, it has been proposed to construct nonparametric shape prior by extending the Parzen density estimator to the space of shapes. For instance, in [5, 20–22], authors proposed a nonlinear statistical shape model for level set segmentation which can be efficiently implemented. Given a set of training shapes, they performed kernel density estimation in the low dimensional subspace. In this way, they are able to combine an accurate model of the statistical shape distribution with efficient optimization in a finite-dimensional subspace. In a Bayesian inference framework, they integrated the nonlinear shape model with a nonparametric intensity model and a set of pose parameters which are estimated in a more direct data-driven manner than in previously proposed level set methods. Kim et al. [14] proposed a nonparametric shape prior model for image segmentation problems. Given example training shapes, they estimate the underlying shape distribution by extending a Parzen density estimator to the space of shapes. Such density estimates are expressed in terms of distances between shapes. The learned shape prior distribution is then incorporated into a maximum a posteriori estimation framework which is solved using active contours.

Recently, Foulonneau et al. [9] proposed an alternative approach for shape prior integration within the framework

of parametric snakes. They combined a compact, parametric representation of shapes within curve evolution theory. More specifically, they proposed to define a *geometric* shape prior based on a description of the target object shape using Legendre moments. A new shape energy term, defined as the distance between moments calculated for the evolving active contour and the moments calculated for a fixed reference shape prior, is proposed and derived in the mathematical framework of [1] in order to obtain the evolution equation. Initially, the method was designed for a single reference shape prior [8], but in the most recent version is able to take into account multi-reference shape priors. As a result, the authors have defined a new efficient method for region-based active contours integrating static shape prior information. Nevertheless, one of the main drawbacks of such an approach lies in its strong dependence to the shape alphabet used as reference. Indeed, as stated by the authors themselves in [9], this method is more related to *template matching* than to *shape learning*.

Inspired by the aforementioned results and especially by the approach proposed by Foulonneau et al., the method proposed in this paper optimizes, within the level sets framework, model consisting of a prior shape probability model and image likelihood function conditioned on shapes. The statistical shape model results from a learning process based on nonparametric estimation of a prior probability, in a low dimensional shape space of Legendre moments built from training silhouette images. Such approach tends to combine most of the advantages of the aforementioned methods, that is to say, it can handle multi-modal shape distributions, preserve a consistent framework for shape modeling and is free from any explicit shape distribution model.

The structure of this paper is as follows: Section 2 describes the proposed image segmentation framework. More specifically in Sect. 2.1 a shape representation using Legendre moments is introduced; The statistical shape model constructed in the space of the Legendre moments is explained in Sect. 2.2; The level set active contour model used in the proposed method is briefly explained in Sect. 2.3; Section 2.4 recasts the energy minimization problem in the general maximum a posteriori (MAP) framework, whereas in Sect. 2.5 the proposed strategy for energy minimization is explained in detail; Section 3 demonstrates the performance of the proposed method on binary silhouette and gray scale images, emphasising the resilience of the proposed method with respect to severe random and structural noise present in the image; Finally the conclusions are given in Sect. 4.

2 Segmentation Framework

The proposed segmentation framework can be seen as constrained contour evolution, with the evolution driven by an

iterative optimization of the posterior probability model that combines a prior shape probability and an image likelihood function linked with a coupling prior imposing constraints on the contour evolution in the image domain. The method can be implemented with any combination of the shape descriptors and dimensionality reduction techniques as long as the shape reconstruction is possible from the selected low dimensional representation. Although for the clarity of the presentation and due to analysis in the experimental section comparing the proposed method against [9], Legendre moments are used in the paper, other shape descriptors such as Zernike moments [23] could be equally used.

In this section all the elements of the proposed model along with the proposed optimization procedure are described in detail.

2.1 Shape Representation Using Legendre Moments

The method proposed in this paper can utilize any shape descriptor as long as it enables shape reconstruction [18, 23]. However, in order to simplify description of the method and comparison with other approaches [2, 9] shapes are encoded, as in [9], by central-normalized Legendre moments $\lambda = \{\lambda_{pq}, p + q \leq N_o\}$ of order N_o where p and q are non-negative integers, and therefore $\lambda \in \mathbf{R}^{N_f}$ with $N_f = (N_o + 1)(N_o + 2)/2$.

The central-normalized Legendre moments are attractive for shape representation as they can be used for objects in arbitrary dimensional spaces and having different topology. They are also invariant to shape scaling and translation and provide compact shape representation where a tradeoff between feature space dimension and shape representation accuracy can be simply controlled by the single parameter N_o . Figure 1 shows an example of shape reconstruction when different values of N_o are used.

For a given shape Ω the moments are defined by:

$$\lambda_{pq} = \frac{1}{|\Omega|} \int_{\Omega} L_{pq}(x, y, \Omega) dx dy \quad (1)$$

where the 2D central-normalized Legendre polynomials L_{pq} are the tensor product of two 1D central-normalized

Legendre polynomials L_p and L_q :

$$L_{pq}(x, y, \Omega) = L_p\left(\frac{x - \bar{x}}{|\Omega|^{1/2}}\right) L_q\left(\frac{y - \bar{y}}{|\Omega|^{1/2}}\right) \quad (2)$$

with Legendre polynomials defined on the interval $[-1, 1]$ as:

$$L_n(x) = \sqrt{\frac{2n+1}{2}} \frac{1}{2^n n!} \frac{d^n}{dx^n} [(x^2 - 1)^n] \quad (3)$$

The area $|\Omega|$ and the center of gravity coordinates (\bar{x}, \bar{y}) are calculated from:

$$|\Omega| = \int_{\Omega} dx dy, \quad (4)$$

$$\bar{x} = \frac{1}{|\Omega|} \int_{\Omega} x dx dy, \quad \bar{y} = \frac{1}{|\Omega|} \int_{\Omega} y dx dy \quad (5)$$

The Legendre polynomials form the orthonormal basis:

$$\int_{-1}^1 L_m(x) L_n(x) dx = \delta_{mn} \quad (6)$$

and therefore are very effective for shape representation. Although the central-normalized Legendre moments provide only scale and translation invariance, the theory presented in this section can be further extended to provide similarity or affine transformation invariance of the moments. Such extension has been well exposed in [9].

In the following sections the scale and translation invariant moments are used but the method would remain the same if similarity or affine invariant moments were used instead.

2.2 Statistical Shape Model of Legendre Moments

In the method proposed here the prior shape constraint is introduced into the segmentation process in the form of probability density function defined on the low dimensional shape space [4] and estimated using Parzen window method. The shape space is constructed using PCA method on a training set consisting of N_s binary silhouette images with foreground and background represented respectively by ones and zeros. The training data can be obtained from previously segmented images or generated from computer models of the objects of interest. In the first instance the central-normalized Legendre moments $\{\lambda_i\}_{i=1}^{N_s}$ are calculated for the shapes $\{\Omega_i\}_{i=1}^{N_s}$ from the training database. Following the methodology proposed in [4] the mean vector $\bar{\lambda}$ and the $N_f \times N_f$ covariance matrix \mathbf{Q} are estimated using:

$$\bar{\lambda} = \frac{1}{N_s} \sum_{i=1}^{N_s} \lambda_i \quad (7)$$

$$\mathbf{Q} = \frac{1}{N_s} \sum_{i=1}^{N_s} (\lambda_i - \bar{\lambda})(\lambda_i - \bar{\lambda})^T \quad (8)$$

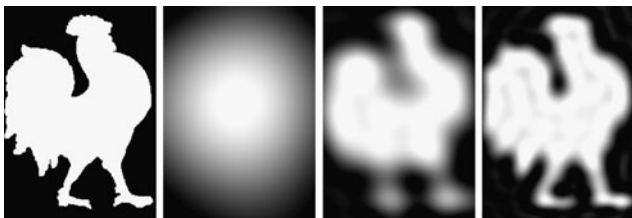


Fig. 1 Images reconstructed from the Legendre moments with different orders. From left to right: original image and reconstruction images with orders $N_o = 5, 20, 40$

Subsequently the $N_f \times N_c$ projection matrix \mathbf{P} is formed by the eigenvectors of the covariance matrix \mathbf{Q} that correspond to the largest N_c ($N_c \leq \min\{N_s, N_f\}$) eigenvalues. The projection of feature vectors $\{\lambda_i\}_{i=1}^{N_s}$ onto the shape space, spanned by the selected eigenvectors, forms the feature vectors $\{\lambda_{r,i}\}_{i=1}^{N_s}$:

$$\lambda_{r,i} = \mathbf{P}^T (\lambda_i - \bar{\lambda}) \quad (9)$$

The density estimation $P(\lambda_r)$, with λ_r defined in the shape space, is performed up to a scale, using $\lambda_{r,i}$ as samples from the population of shapes and with the isotropic Gaussian function as the Parzen window:

$$P(\lambda_r) = \sum_{i=1}^{N_s} \mathcal{N}(\lambda_r; \lambda_{r,i}, \sigma^2) \quad (10)$$

where $\mathcal{N}(\lambda_r; \lambda_{r,i}, \sigma^2) = \exp(-\|\lambda_r - \lambda_{r,i}\|^2 / 2\sigma^2)$

2.3 Level Set Active Contour Model

Introduced in the previous section, density function $P(\lambda_r)$ is defined on the shape space of Legendre moments and represents a prior knowledge learned from the training shape examples.

To detect and segment shapes present in an observed image, a mechanism for taking into consideration the evidence about shape needs to be included. Due to the way the final objective function is optimized, any energy-based level-set contour evolution schemes can be used. In this paper, it is proposed to consider for this purpose active contours implemented in the level set framework. The region competition scheme proposed by [2] will be used for the illustration purposes. In this case, it is assumed that the image I is formed by regions of approximatively constant intensity values and the segmentation is defined as energy minimization problem, with the energy given by:

$$E_{cv}(\Omega, \mu_\Omega, \mu_{\Omega^c} | I) = \int_{\Omega} (I - \mu_\Omega)^2 dx dy + \int_{\Omega^c} (I - \mu_{\Omega^c})^2 dx dy + \gamma |\partial\Omega| \quad (11)$$

where Ω^c represents the complement of Ω in the image domain and $|\partial\Omega|$ represents the length of the boundary $\partial\Omega$ of the region Ω . The above defined energy minimization problem can be equivalently expressed as maximization of the likelihood function:

$$P(I|\Omega) \propto \exp(-E_{cv}(\Omega, \mu_\Omega, \mu_{\Omega^c} | I)) \quad (12)$$

where $P(I|\Omega)$ could also be interpreted as a probability of observing image I when shape Ω is assumed to be present

in the image. Introducing level set (embedding) function ϕ such that the Ω can be expressed in terms of ϕ as $\Omega = \{(x, y) : \phi(x, y) \geq 0\}$, as well as $\Omega^c = \{(x, y) : \phi(x, y) < 0\}$ and $\partial\Omega = \{(x, y) : \phi(x, y) = 0\}$, the foregoing functional is equivalent to

$$E_{cv}(\phi, \mu_\Omega, \mu_{\Omega^c} | I) = \int (I - \mu_\Omega)^2 H(\phi) dx dy + \int (I - \mu_{\Omega^c})^2 (1 - H(\phi)) dx dy + \gamma \int |\nabla H(\phi)| dx dy \quad (13)$$

with H representing Heaviside function. Calculating Gateaux derivative [1] it can be shown that such energy function is minimized by function ϕ given as a solution of the following PDE equation

$$\frac{\partial\phi}{\partial t} = \left((I - \mu_{\Omega^c})^2 - (I - \mu_\Omega)^2 \right) |\nabla\phi| + \gamma \nabla \left(\frac{\nabla\phi}{|\nabla\phi|} \right) \cdot \nabla\phi \quad (14)$$

with $\mu_\Omega = \int_{\Omega} I dx dy$ and $\mu_{\Omega^c} = \int_{\Omega^c} I dx dy$ representing respectively the average intensities inside and outside the evolving curve.

2.4 MAP Framework

Introduced in the previous two sections, distributions representing shape prior information and image intensity can be combined using Bayes rule:

$$P(\lambda_r | I) \propto P(\lambda_r) P(I | \lambda_r) \quad (15)$$

where $P(\lambda_r)$ and $P(I | \lambda_r)$ represent respectively shape and intensity based information. In [26] it was proposed to optimize $P(\lambda_r | I)$ by restricting the shape evolution in the estimated shape space, by imposing following constraint: $P(I | \lambda_r) = P(I | \Omega) |_{\Omega=\Omega(\lambda_r)}$. As maximizing $P(\lambda_r | I)$ is equivalent to minimizing $-\ln(P(\lambda_r | I))$, Zhang et al. [26] suggested minimizing an energy function:

$$E(\lambda_r) = E_{prior}(\lambda_r) + E_{image}(\lambda_r) \quad (16)$$

where the shape prior term is defined as:

$$E_{prior}(\lambda_r) = -\ln \left(\sum_{i=1}^{N_s} \mathcal{N}(\lambda_r; \lambda_{r,i}, \sigma^2) \right) \quad (17)$$

and is built based on the shape samples Ω_i as explained in Sect. 2.2. The image term is defined as:

$$E_{image}(\lambda_r) = E_{cv}(\Omega, \mu_\Omega, \mu_{\Omega^c} | I) |_{\Omega=\Omega(\lambda_r)} \quad (18)$$

where optimization of E_{cv} is constraint to shapes Ω from the estimated shape space $\Omega = \Omega(\lambda_r)$ ($\Omega(\lambda_r)$ denotes a shape from the shape space represented by the Legendre moments $\lambda = \mathbf{P}\lambda_r + \bar{\lambda}$).

As it was indicated in [26] such approach provides a very robust segmentation. Unfortunately the solution which minimizes $E(\lambda_r)$ belongs to the shape space and as such may not accurately represent object of interest. To resolve this the Eq. (15) can be redefined as:

$$P(\Omega, \lambda_r | I) \propto P(\lambda_r) P(\Omega | \lambda_r) P(I | \Omega, \lambda_r) \quad (19)$$

Equation (19) is now optimized jointly with respect to shape Ω defined in the image space and vector λ_r defined in the shape space. The coupling between these two is achieved by $P(\Omega | \lambda_r)$ defined as:

$$P(\Omega | \lambda_r) \propto \exp(-E_c(\Omega | \lambda_r)) \quad (20)$$

with:

$$E_c(\Omega | \lambda_r) = \alpha \int (H(\phi) - H(\phi_r))^2 dx dy \quad (21)$$

where α is a weighting factor defining the strength of coupling between Ω and ϕ_r is a signed distance function representing the shape defined by the λ_r in the image domain. The overall energy to be minimized is now given by:

$$E(\Omega, \lambda_r) = E_{prior}(\lambda_r) + E_{image}(\Omega, \lambda_r) \quad (22)$$

with the image energy:

$$E_{image}(\Omega, \lambda_r) = E_{cv}(\Omega, \mu_\Omega, \mu_{\Omega^c} | I) + E_c(\Omega | \lambda_r) \quad (23)$$

It can be shown that the corresponding PDE describing the solution of this new image energy is given by:

$$\begin{aligned} \frac{\partial \phi}{\partial t} = & \left((I - \mu_{\Omega^c})^2 - (I - \mu_\Omega)^2 \right) |\nabla \phi| \\ & + \alpha (2H(\phi_r) - 1) |\nabla \phi| + \gamma \nabla \left(\frac{\nabla \phi}{|\nabla \phi|} \right) |\nabla \phi| \end{aligned} \quad (24)$$

The details of the optimization procedure for energy $E(\Omega, \lambda_r)$ are given in the next section.

2.5 Optimization

In the implementation of the proposed method the energy given in Eq. (22) is minimized using a greedy method where each of the two energy components E_{prior} and E_{image} is minimized in turn. The optimization of the image based energy E_{image} is implemented through evolution of the level set ϕ defined by Eq. (24) with λ_r fixed. Subsequently the E_{prior} is minimized in the shape space with respect to the λ_r . In this approach active contour evolution can be interpreted as a method for transferring the evidence about the

shape present in the image into the shape space where it is combined with the shape information derived from the training shape samples.

The overall optimization procedure is summarized in the following steps:

- Projection of the current shape $\Omega^{(k)}$ into the shape space:

$$\Omega^{(k)} \rightarrow \lambda_r^{(k)} \quad (25)$$

where $\lambda_r^{(k)} = \mathbf{P}^T(\lambda^{(k)} - \bar{\lambda})$, and the central-normalized Legendre moments in vector $\lambda^{(k)}$ are calculated using:

$$\lambda_{pq}^{(k)} = \frac{1}{|\Omega^{(k)}|} \int_{\Omega^{(k)}} L_{pq}(x, y, \Omega^{(k)}) dx dy \quad (26)$$

where $\Omega^{(k)}$, comes from the previous algorithm iteration;

- Shape space vector update:

$$\lambda_r^{(k)} \rightarrow \lambda_r'^{(k)} \quad (27)$$

This step reduces the value of E_{prior} by moving $\lambda_r^{(k)}$ in the steepest descent direction:

$$\lambda_r'^{(k)} = \lambda_r^{(k)} - \beta \left. \frac{\partial E_{prior}}{\partial \lambda_r} \right|_{\lambda_r = \lambda_r^{(k)}} \quad (28)$$

where

$$\frac{\partial E_{prior}}{\partial \lambda_r} = \frac{1}{2\sigma^2} \sum_{i=1}^{N_s} w_i (\lambda_r - \lambda_{r,i}) \quad (29)$$

with

$$w_i = \frac{\mathcal{N}(\lambda_r; \lambda_{r,i}, \sigma^2)}{\sum_{k=1}^{N_s} \mathcal{N}(\lambda_r; \lambda_{r,k}, \sigma^2)} \quad (30)$$

- Shape reconstruction from Legendre moments:

$$\lambda_r'^{(k)} \rightarrow \Omega'^{(k)} \quad (31)$$

where shape $\Omega'^{(k)}$ is reconstructed using:

$$\begin{aligned} \Omega'^{(k)} = & \left\{ (x, y) : \left(\sum_{p,q}^{p+q \leq N_o} \lambda_{pq}'^{(k)} L_{pq}(x, y, \Omega'^{(k)}) \right) > 0.5 \right\} \end{aligned} \quad (32)$$

with the Legendre moments $\lambda_{pq}'^{(k)}$ in vector $\lambda'^{(k)}$ calculated from the shape space vector $\lambda_r'^{(k)}$ using: $\lambda'^{(k)} = \mathbf{P}\lambda_r'^{(k)} + \bar{\lambda}$

- Evolution of $\Omega'^{(k)}$ according to Eq. (24):

$$\Omega'^{(k)} \rightarrow \Omega^{(k+1)} \quad (33)$$

shape $\Omega'^{(k)}$, is a shape represented in the shape space and $\Omega^{(k+1)}$ is the result of shape evolution in the image domain;

These steps are iterated until no shape change occurs in two consecutive iterations: $\Omega^{(k+1)} = \Omega^{(k)}$.

The proposed strategy provides the maximum flexibility by making the optimizations in image space and shape space two independent processes bridged by shape projection and reconstruction. Thus, changing the curve evolution model in the image space or probability estimation model in the shape space will not affect other procedures. Although Legendre moments and PCA are selected to build the shape space in this paper, other shape descriptors and dimensionality reduction techniques can be easily ‘plugged’ into the optimization framework as long as the shape reconstruction from the shape space is possible. It should be pointed out that, unlike derivative based optimization methods such as [8] and [9], the shape descriptors need *not* be differentiable in the proposed method.

To guarantee convergence of the algorithm the parameter α in equation Eq. (24) should be non-decreasing function of the iteration index. In that case the convergence is guaranteed as for large enough value of α the algorithm, if not terminated beforehand, is equivalent to the steepest descent in the reduced shape space. In practical implementation the value of α is periodically increased after predefined number of iterations lapses. With this in mind the proposed algorithm can be interpreted as a mode seeking shape detection procedure. With small value of α the algorithm can relatively easily make long “unconstrained” jumps in the shape space following the shape evidence in the image domain. With the gradually increasing value of α the algorithm will be restricted to make gradually smaller steps to maintain similarity of the evolving shape in the image domain to the current shape defined in the shape space. It should be noted that in the practical experiments the algorithm converged in just a few iterations without increasing α for the vast majority of cases. To further improve segmentation results after the algorithm terminates the image energy can be minimized independently through the contour propagation defined by formula Eq. (24). In this case the value of the parameter α should correspond to the level of noise present in the image, with small values of α corresponding to low level of noise. This is further explained in the experimental section.

3 Experimental Results

To evaluate the proposed method, experiments were carried out using binary silhouette and real gray scale images. The main reason behind using the silhouette images was to investigate robustness of the proposed technique against severe random and structural noise present in data. The segmentation of such images without any noise is straightforward, as it could be achieved by simple thresholding, proving ready ground truth data. Additionally any incorrect segmentation

of the noisy images can be directly associated with the noise rather than with a specific “non-optimal” type of image intensity descriptor used to compute the external energy in the active contour model. As it was explained in Sect. 2 the proposed method can be used with any contour evolution equation and as such can be used with colour or even tensor valued data. Here for illustration purposes results showing segmentation of real gray scale images were included.

3.1 Silhouette Data

A first set of experiments were carried out using a chicken image set consisting of 20 binary silhouette images of different shapes, orientations and sizes from the *MPEG7 CE shape-1 Part B* database [15]. The first 19 of them were used as training shapes for building the statistical prior model and the remaining image was used for testing (see Fig. 2). The diversity of the training shapes can be clearly noted—rotations in the images were not removed on purpose to test robustness of the proposed method against large shape variability.

Shape variations represented by the sample points along the three most dominant principal axes in the feature space are shown in Fig. 3, from which it can be observed that the three principal axes respectively capture the shape variabilities of rotation (vertical position v.s. horizontal position), trend (V-shape v.s. L-shape) and reflection (right headed v.s. left headed).

The selected chicken’s silhouette image (Fig. 4(a)) was corrupted by a combination of two different types of noise, namely, the additive white Gaussian noise for the simulation of a sensor noise and a structural noise simulating occlusions and defects.

The test image with Gaussian noise is shown in Fig. 4(b) where the noise level is so high that even with a prior knowledge of the shape it is difficult to find the original silhouette

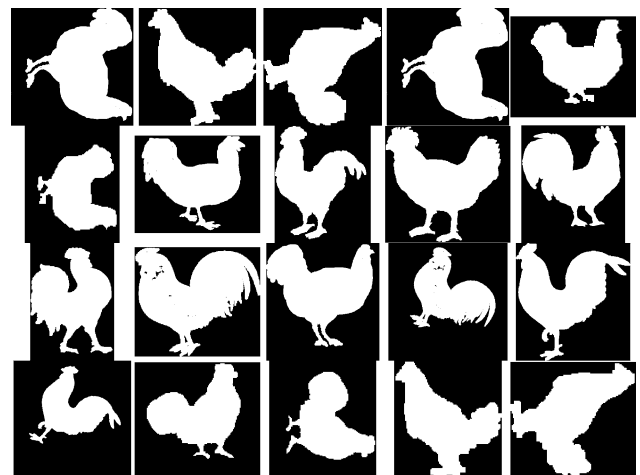


Fig. 2 The chicken image set

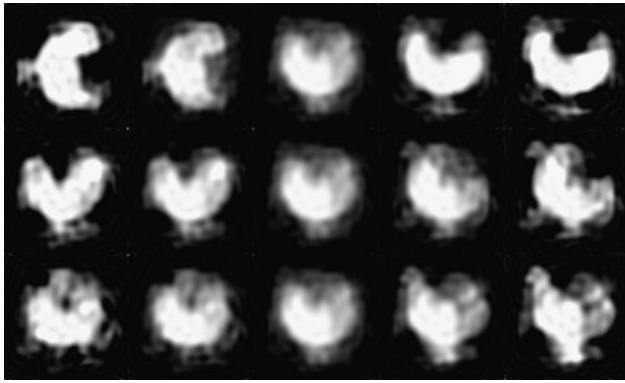


Fig. 3 PCA results on the chicken images. From *top* to *bottom*, the three rows represent the shapes reconstructed from the Legendre moments sampled along the three most dominant principal axes (eigenvectors) in the feature space. From *left* to *right*, each column corresponds to different magnitude (-2 , -1 , 0 , 1 , 2 times the squared root of the eigenvalue associated with the corresponding eigenvector) of shape variations from the mean shape

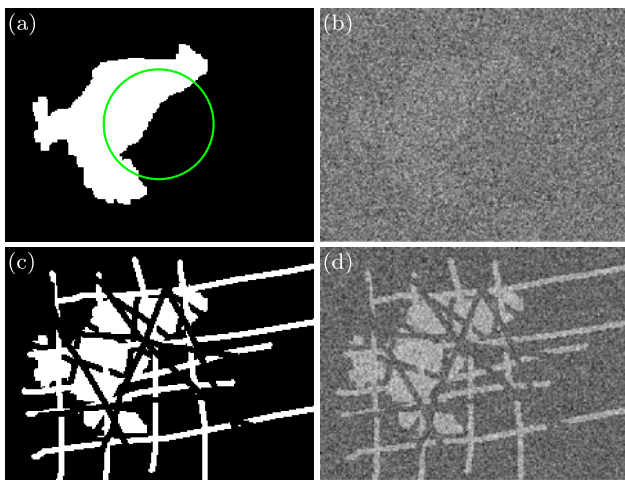


Fig. 4 Test images, (a) original binary silhouette image with initial active contour used in the experiments, shown as a *circle* at the center of the image, (b) test image corrupted by Gaussian noise, (c) test image with structural noise, (d) test image with hybrid (Gaussian and structural) noise

in the noisy image. For the structural noise (Fig. 4(c)), hard alterations are made on the original image in order to emphasize the need for shape constraints. Finally, the last test image is corrupted by both Gaussian and structural noise (Fig. 4(d)).

All the experimental results shown in Fig. 5 were based on the test images as shown in Fig. 4 and with the same parameters $N_o = 40$ and $N_c = 10$, used to calculate Legendre moments and the shape space.

As it can be seen in Fig. 5 for each corrupted image, the proposed method makes a satisfactory shape segmentation even though this shape was not included in the database of the shapes used to calculate the shape space. Figure 5(d)

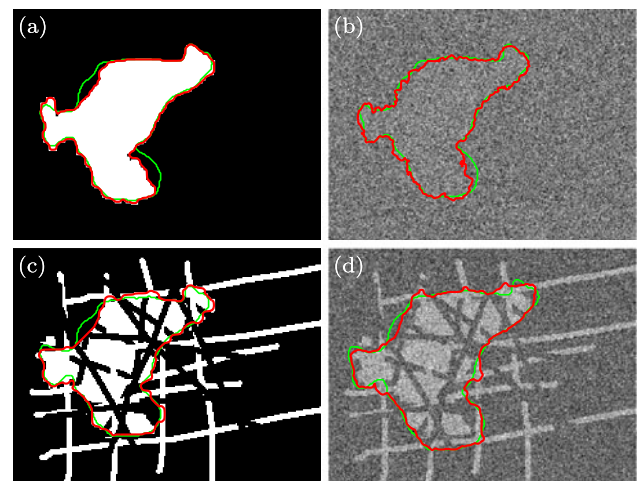


Fig. 5 Results obtained by the proposed method for: (a) noise-free test image; (b) test image corrupted by the Gaussian noise; (c) test image with the structural noise; (d) test image corrupted by the hybrid noise. Whereas the *green* line represents the solutions defined in the shape space corresponding to $\alpha = \infty$ in the final algorithm iteration, the *red* line shows solutions obtained for α chosen based on the level of noise present in the images, e.g. for noise free image shown in (a) $\alpha = 0$

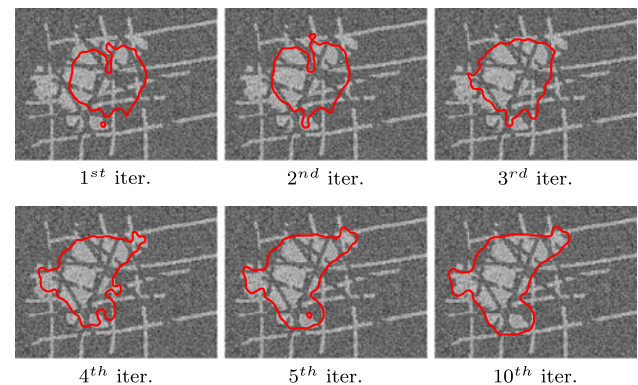


Fig. 6 Intermediate contour evolution in the shape space obtained for the result shown in Fig. 5(d)

shows the final segmentation results for the image corrupted by both Gaussian and structural noise. It can be seen that in the solution, following Eq. (15), defined in the shape space (shown in green) detailed shape variabilities are normally missing. This is most prominent in the noiseless image. Whereas the solution defined in the image space, Eq. (19), the corresponding Ω (shown in red), closely follows edges of the silhouette. For the noisy images, particularly with the severe random noise, quality of the segmentation in the image domain slightly deteriorates. This can be understood as manifestation a basic tradeoff between fidelity and robustness to noise. In the proposed method this tradeoff is controlled by the α parameter, Eq. (24), where small value of α encourages fidelity whereas larger values improve robustness of the solution. Samples of the evolving shape for this specific test image are shown in Fig. 6.

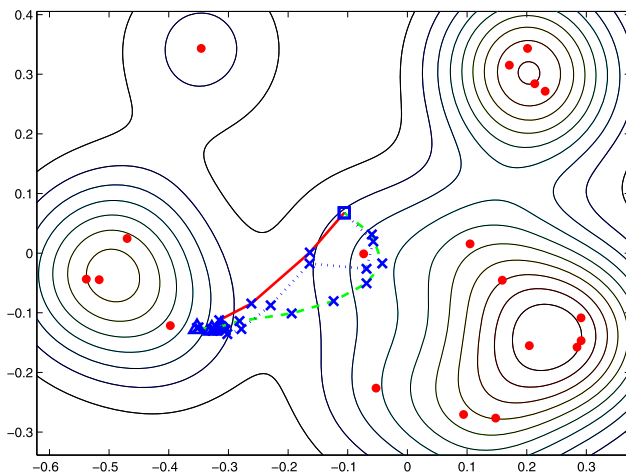


Fig. 7 Shape evolution trajectories shown in the feature space spanned by the first two principal axes

Figure 7 shows the shape evolution trajectories in the feature space, spanned by the first two principal axes, corresponding to the results shown in Fig. 5. The iso-contours shown in this figure illustrate the probability density function (pdf) estimated using isotropic Gaussian function as the Parzen window with $\sigma^2 = 0.02$. The dots represent the projections of the 19 training shapes.

The three curves in Fig. 7, shown in solid, dash and dotted lines, respectively demonstrate the trajectories formed by the optimization processes of the proposed method based on the test images with Gaussian, structural and hybrid noise. As the same initial circular shape was used for all three test images all the trajectories start at the same point marked by a square. All trajectories converge to points scattered nearby the dot representing the shape included in the image shown in the first row and third column in Fig. 2, which is the most similar to the shape present in the test images. Focusing on the dotted trajectory within the feature space, one can match trajectory steps with the intermediate results shown in Fig. 6. The fact that the convergent points are close but not exactly on the dot indicates that the proposed approach is not a template matching. Although the method is designed to search for shapes similar to shapes seen during the training process it can recover some unseen shape variations.

To assess the performance of the proposed method, the results obtained using proposed method were compared against the segmentation results generated by the Chan-Vese model (without shape constraint) shown in Fig. 8 and with the result obtained using the multi-reference method proposed in [9] shown in Fig. 9.

The segmentation result for the additive Gaussian noise from the Chan-Vese model, which is well-known for its robustness to Gaussian noise, is shown in Fig. 8(b). Inaccurate as it is, the result does provide some reasonable indications about the shape and position of the desired object, shown as

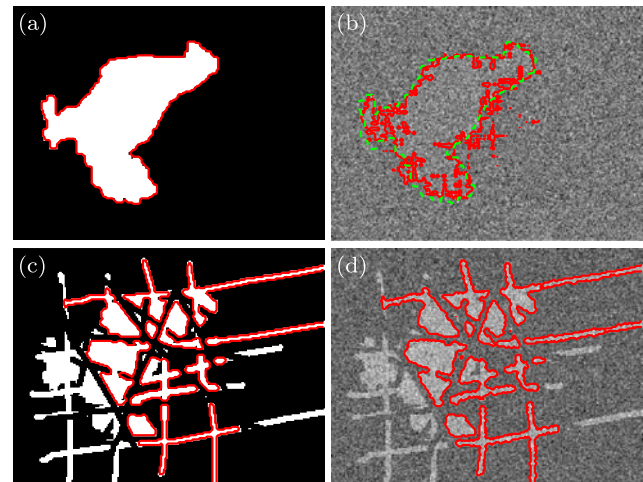


Fig. 8 Segmentation results obtained for the corresponding test images from Fig. 4 using Chan-Vese model

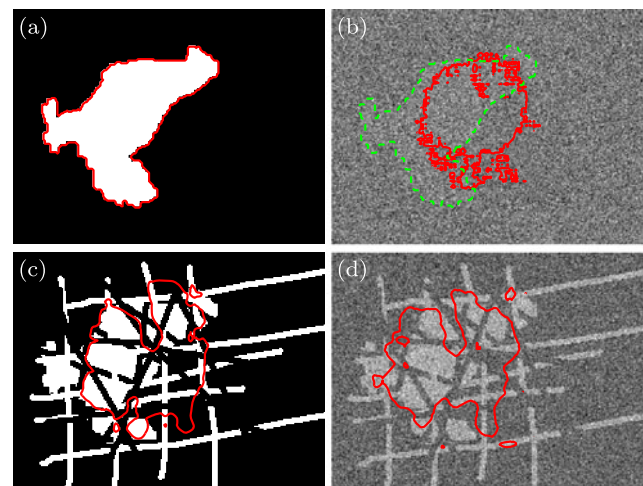


Fig. 9 Segmentation results obtained for the corresponding test images from Fig. 4 using multi-reference method proposed in [9]

a dash line, which is one of the major reasons why region-based active contour approaches such as Chan-Vese model are good choices for the image term in the proposed method. Figure 9(b) shows the segmentation result using the multi-reference method from [9], where all the 20 training shapes were used as references. The result demonstrates a dilemma for the methods with ‘soft’ shape constraints—How to or is it possible to select an appropriate weight to balance the image term and shape term? For a noisy image like this, a strong image force could lead to the inaccurate result as shown in Fig. 8(b), whereas a strong shape force could result in the convergence to a wrong shape at a wrong location due to the lack of guidance from image force. In this case, a range of different weights were tried, but none of them converged to the right result. Much better result was achieved using the proposed method as shown in Fig. 5(b).

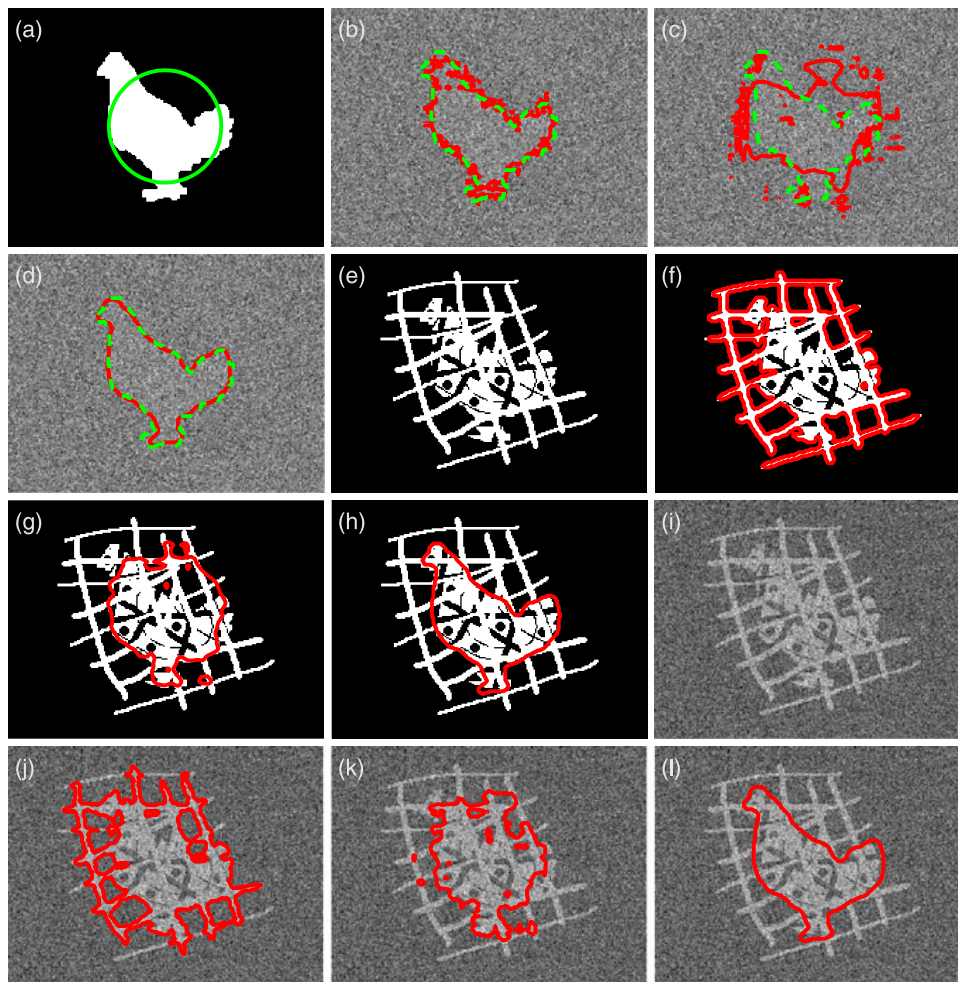


Fig. 10 Results for a second set of experiments using different chicken's silhouette image. This specific image has been removed from the database before building the shape space subsequently used in the iterations. The segmentation results are shown as *red solid curves*, whereas the desired results are shown as *green dash lines*. (a) Original noise-free test image with initial active contour shown as a *circle* at the *center* of the image; (b) Segmentation of the test image with severe Gaussian noise using Chan-Vese method; (c) Segmentation of

the same test image as in (b) using the multi-reference method proposed in [9]; (d) Segmentation of the same test image as in (b) using the proposed method; (e) Test image with structural noise; (f) Segmentation of (e) using Chan-Vese model; (g) Segmentation of (e) using the multi-reference method; (h) Segmentation of (e) using the proposed method; (i) Test image with hybrid noise; (j) Segmentation of (i) using Chan-Vese method; (k) Segmentation of (i) using the multi-reference method; (l) Segmentation of (i) using the proposed method

As expected, the resulting shape living in the reduced feature space tends to have more regular appearance.

For images with a large amount of structural noise Chan-Vese model without shape constraint completely failed, as shown in Fig. 8(c–d), by following the false structures. Although increasing the weight associated with the length term (γ in Eq. (11)) can avoid some of the false structures, it cannot properly locate the desired shape. Again, the multi-reference method failed to converge to the right result as evident from Fig. 9(c–d).

Figure 10 collocates the results obtained for a different test images generated from the different chicken silhouette. As before the selected image was removed from the training set prior to construction of the shape space.

Once again, the proposed method leads to the most satisfying results. Figure 11 demonstrates the trajectories formed by the optimization processes of the proposed method applied to the data with Gaussian, structural and hybrid noise. It can be noticed that the local pdf maxima are “better defined” in comparison to the pdf shown in Fig. 7 as in this case a smaller value of $\sigma^2 = 0.002$ was used within the Gaussian kernel. Regarding convergence of the different trajectories, the same conclusions as in the first set of experiments can be made.

Although the main objective of the described experiment was to demonstrate a superior robustness of the proposed methods with respect to severe random and structural noise, the accuracy of the method was also tested on repeated ex-

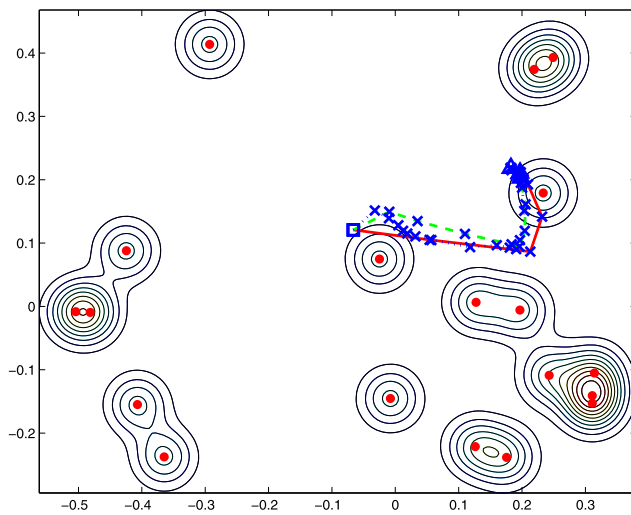


Fig. 11 Trajectories in the first two principal axes of the shape space: (i) *solid* for data with Gaussian noise; (ii) *dash* for data with structural noise; (iii) *dotted* for data with hybrid noise



Fig. 12 Training set used to build the shape space for the cup object

periments with different combination of the target image and structural noise pattern. It transpired that the proposed method was able to localize object boundary with an average accuracy of 1.2, 1.7 and 2 pixels when operating respectively on images with Gaussian, structural and hybrid noise.

3.2 Gray Scale Images

Finally experiments were carried out using a gray scale images to test performance of the proposed methodology on real images. The first test image used in these experiments is shown in Fig. 13(a) where the objective was to segment the cup. The shape space was constructed from the image set shown in Fig. 12, with a subset of the *MPEG7 CE shape-1 Part B* database used. It can be clearly seen that the training shapes integrate a large shape variability, and that differ-

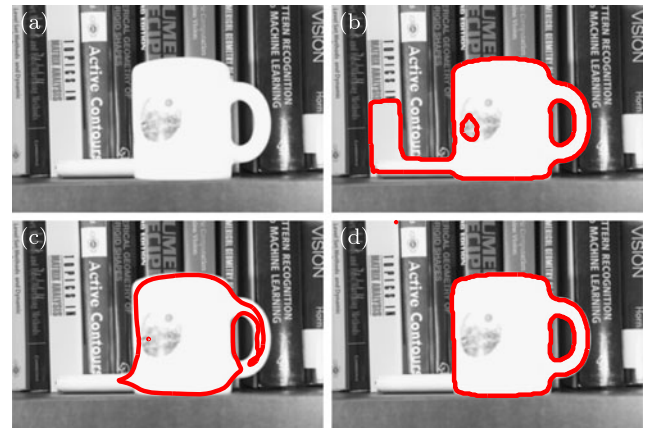


Fig. 13 Segmentation results for the cup image (a) an image to be segmented, (b) result of segmentation using Chan-Vese model, (c) result of the segmentation using the multi-reference method from [9], (d) result of the segmentation using the proposed method

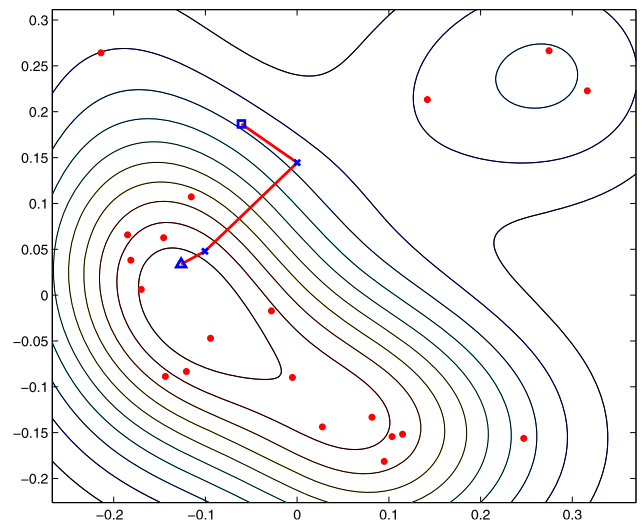


Fig. 14 Trajectory in the space of the first two principal directions of the cup shape space corresponding to result shown in Fig. 13. The *square* represents the starting point, the *triangle* the projection of the final detected shape

ent positions of the handle are taken into account (left and right). Results of segmentation using the Chan-Vese, multi-reference and the proposed method are shown in Fig. 13.

Assuming that the goal of the segmentation was to recover the shape of the cup, the proposed method leads to more accurate result with the final shape segmentation not altered by the drawing on the cup or by books and a pen in the background. The corresponding trajectory of the optimization process can be seen in Fig. 14.

This demonstrates that, the proposed method is more robust than the other two tested methods with respect to “shape distractions” present in the data. The final result can be seen as a good compromise between image information and the prior shape constraints imposed by the training data set used.

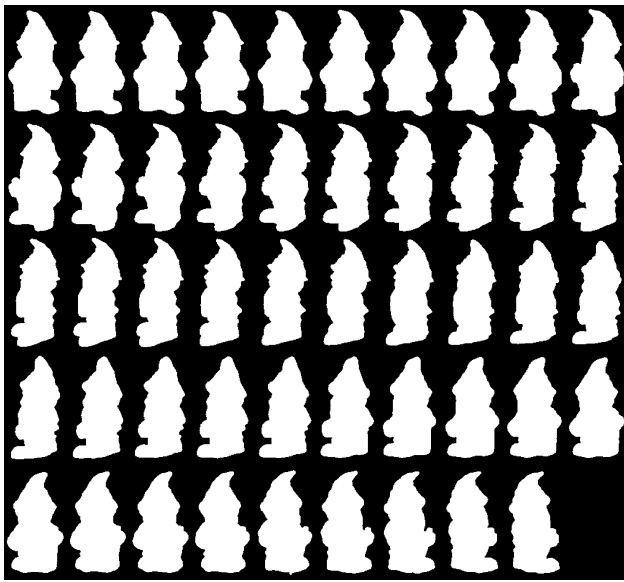


Fig. 15 Training set used to build the shape space for the garden gnome figurine

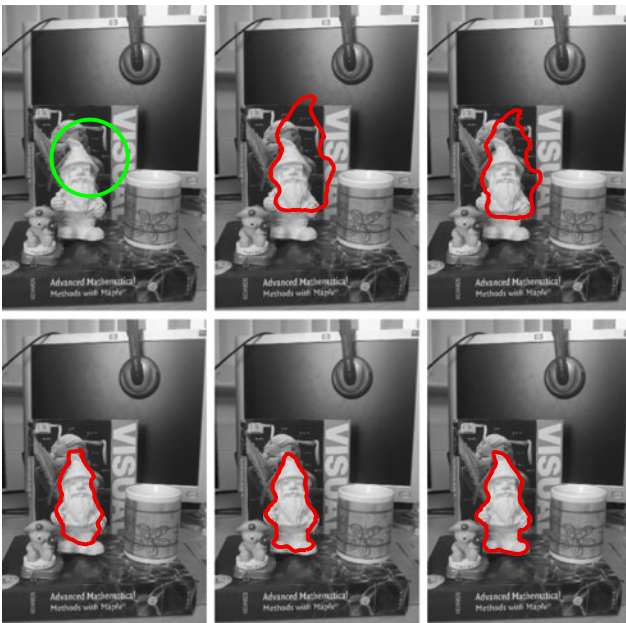


Fig. 16 Contour evolution for the garden gnome figurine showing results for iterations 0 (initial contour), 5, 15, 25, 30 and 39 (stable contour)

The second test image used in the experiments is shown in Fig. 16, where the objective is to segment the garden gnome figurine. This time the training data was obtained from a video sequence showing the same figurine rotating against a black background. The silhouettes used for training are shown in Fig. 15. These were extracted from 49 frames, sampled from a video sequence covering approximately 320° of the rotational viewing angle of the figurine.

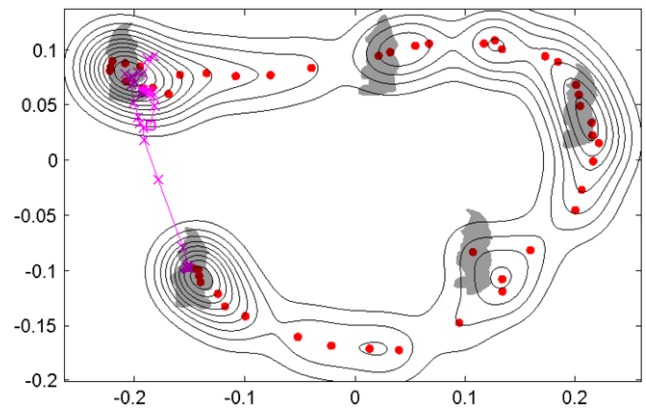


Fig. 17 Shape trajectory shown in the feature space spanned by first two principal directions, corresponding to contour evolution shown in Fig. 16. Superimposed silhouettes show shapes represented at different locations of the shape space

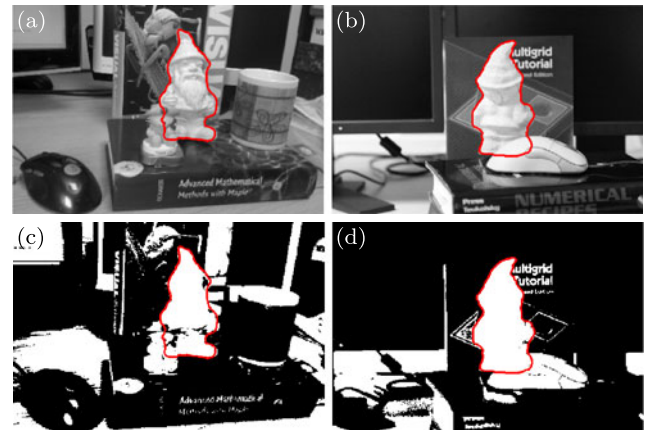


Fig. 18 More results obtained for the garden gnome figurine: (a–b) stable contours; (c–d) corresponding segmented propagation fields PF indicating image driven shape internal (white) and external (black) regions

Figure 16 shows initial contour (top left) and the stable contour (bottom right) together with a sample of the intermediate shapes demonstrating the contour convergence process.

Figure 17 shows corresponding shape evolution in the feature space. It could be noticed that although contour was initially evolving around shapes shown in the bottom row of Fig. 15, represented in the top left part of the feature space, it was eventually able to cross a “long valley” in the feature space and converge to a shape in the vicinity of the shapes shown in the top row of Fig. 15.

Figure 18 shows results obtained for two other images depicting the same garden gnome figurine. In both cases the contours converged to the correct shapes despite the occlusions and the distractors in the background. To demonstrate the robustness of the method the corresponding segmented propagation fields $PF = ((I - \mu_{\Omega^c})^2 - (I - \mu_{\Omega})^2)$, calcu-

lated in Eq. (24) as part of contour update in the image domain, are shown in Fig. 18(c–d). In those images white and black pixels correspond respectively to the positive and negative values of the PF function. It can be clearly noticed that the shape priors are activated in both cases as the stable contours ignore erroneous image driven shape internal/external region indicators.

4 Conclusions

The paper describes a novel method for shape detection and image segmentation. The proposed method can be seen as constrained contour evolution, with the evolution driven by an iterative optimization of the posterior probability function that combines a prior shape probability, the coupling distribution, and the image likelihood function. The prior shape probability function is defined on the subspace of Legendre moments and is estimated, using Parzen window method, on the training shape samples given in the estimated beforehand shape space. The likelihood function is constructed from conditional image probability distribution, with the image modeled to have regions of approximately constant intensities. The coupling distribution is defined as the prior distribution on the image likelihood function which imposes feasible shapes changes based on the current shape parametrization in the shape space. The resulting constrained optimization problem is solved using combinations of level set active contour evolution in the image space and steepest descent iterations in the shape space. The decoupling of the optimization processes into image and shape spaces provides an extremely flexible optimization framework for general statistical shape based active contour where evolution function, statistical model, shape representation all become configurable. The presented experimental results demonstrate very strong resilience of the proposed method to the random as well as structural noise present in the image.

References

1. Aubert, G., Barlaud, M., Faugeras, O., Jehan-Besson, S.: Image segmentation using active contours: calculus of variations or shape gradients? *SIAM J. Appl. Math.* **63**, 2128–2154 (2003)
2. Chan, T.F., Vese, L.A.: Active contours without edges. *IEEE Trans. Image Process.* **10**(2), 266–277 (2001)
3. Chen, Y., Tagare, H., Thiruvankadam, S., Huang, F., Wilson, D., Gopinath, K., Briggs, R., Geiser, E.: Using prior shapes in geometric active contours in a variational framework. *Int. J. Comput. Vis.* **50**(3), 315–328 (2002)
4. Cootes, T.F., Taylor, C.J., Cooper, D.H., Graham, J.: Active shape models—their training and application. *Comput. Vis. Image Underst.* **61**(1), 38–59 (1995)
5. Cremers, D., Osher, S.J., Soatto, S.: Kernel density estimation and intrinsic alignment for shape priors in level set segmentation. *Int. J. Comput. Vis.* **69**(3), 335–351 (2006)
6. Erdem, E., Tari, S., Vese, L.: Segmentation using the edge strength function as a shape prior within a local deformation model. In: *ICIP*, pp. 2989–2992 (2009)
7. Ettinger, P., Segonne, F., Keriven, R.: Shape prior using manifold learning techniques. In: *ICCV*, pp. 1–8 (2007)
8. Foulonneau, A., Charbonnier, P., Heitz, F.: Geometric shape priors for region-based active contours. In: *ICIP*, vol. 3, pp. 413–416 (2003)
9. Foulonneau, A., Charbonnier, P., Heitz, F.: Multi-reference shape priors for active contours. *Int. J. Comput. Vis.* **81**(1), 68–81 (2009)
10. Fussenegger, M., Roth, P., Bischof, H., Deriche, R., Pinz, A.: A level set framework using a new incremental, robust Active Shape Model for object segmentation and tracking. *Image Vis. Comput.* **27**(8), 1157–1168 (2009)
11. Gastaud, M., Barlaud, M., Aubert, G.: Combining shape prior and statistical features for active contour segmentation. *IEEE Trans. Circuits Syst. Video Technol.* **14**(5), 726–734 (2004)
12. Houhou, N., Lemkaddem, A., Duay, V., Allal, A., Thiran, J.-P.: Shape prior based on statistical MAP for active contour segmentation. In: *ICIP*, pp. 2284–2287 (2008)
13. Kass, M., Witkin, A., Terzopoulos, D.: Snakes: active contour models. *Int. J. Comput. Vis.* **1**(4), 321–331 (1988)
14. Kim, J., Çetin, M., Willsky, A.S.: Nonparametric shape priors for active contour-based image segmentation. *Signal Process.* **87**(12), 3021–3044 (2007)
15. Latecki, L.J., Lakamper, R., Eckhardt, T.: Shape descriptors for non-rigid shapes with a single closed contour. In: *CVPR*, pp. 424–429 (2000)
16. Lecellier, F., Jehan-Besson, S., Fadili, J., Aubert, G., Revenu, M., Saloux, E.: Region-based active contour with noise and shape priors. In: *ICIP*, pp. 1649–1652 (2006)
17. Leventon, M., Grimson, W., Faugeras, O.: Statistical shape influence in geodesic active contours. In: *CVPR*, pp. 316–323 (2000)
18. Mukundan, R., Ramakrishnan, K.: Moments functions in image analysis-theory and applications. World Scientific, Singapore (1998). ISBN 981-02-3524-0
19. Prisacariu, A.V., Reid, I.: Nonlinear shape manifolds as shape priors in level set segmentation and tracking. In: *CVPR*, pp. 2185–2192 (2011)
20. Rousson, M., Cremers, D.: Efficient kernel density estimation of shape and intensity priors for level set segmentation. In: *MICCAI*, pp. 757–764 (2005)
21. Rousson, M., Paragios, N.: Shape priors for level set representations. In: *ECCV*, pp. 78–92 (2002)
22. Rousson, M., Paragios, N.: Prior knowledge, level set representations and visual grouping. *Int. J. Comput. Vis.* **76**(3), 231–243 (2008)
23. Teague, M.: Image analysis via the general theory of moments. *J. Opt. Soc. Am.* **70**(8), 920–930 (1980)
24. Thiruvankadam, T.R., Chan, T., Hong, B.-W.: Segmentation under occlusions using selective shape prior. *SIAM J. Imaging Sci.* **1**(1), 115–142 (2008)
25. Tsai, A., Yezzi, A., Wells, W.M., Tempany, C., Tucker, D., Fan, A., Grimson, W., Willsky, A.: A shape-based approach to the segmentation of medical imagery using level sets. *IEEE Trans. Med. Imaging* **22**(2), 137–154 (2003)
26. Zhang, Y., Matuszewski, B.J., Histace, A., Precioso, F.: Statistical shape model of Legendre moments with active contour evolution. In: *CAIP*, pp. 51–58 (2011)



Yan Zhang acquired his PhD degree at School of Computing, Engineering and Physical Sciences, University of Central Lancashire, UK in 2006. After that, he worked in the same school as a post-doctoral researcher focusing on medical image processing and analysis which led to more than 10 publications. He's currently working in Shenzhen Landun Technology Co. Ltd., P.R.China, as a Manager of the Machine Vision Department. His current interests include applying computer vision techniques to intelligent transport systems.



Bogdan J. Matuszewski currently holds position of Reader in Computer Vision and heads the Robotics and Vision Research Laboratory at the University of Central Lancashire, Preston, UK. He received his PhD degree in Electronic Engineering from the Wrocław University of Technology, Poland, in 1996. His research interests include: use of Bayesian methodology for modelling, tracking and pattern recognition; deformable models, variational and partial differential equations methods for image analysis

applied to data registration, segmentation and interpretation; biomedical and industrial applications of computer vision and machine learning. He published over 100 research papers in different areas of computer vision and image processing, and successfully supervised 12 PhD students.



Pontoise, Cergy, France with ETIS laboratory (UMR CNRS 8051).

Aymeric Histace was born in 1977 in Lyon, France. He received the Master degree in Signals/Images for Computer Aided Diagnosis in 2001 (University of Angers, France). He received, in 2004, the Ph.D degree in signal and image processing at the Laboratoire d'Ingénierie des Systèmes Automatisés (LISA), University d'Angers, France. He is since 2006, an associate professor in signal and image processing with application to life sciences (mainly Computer Aided Diagnosis) at the University of Cergy-



Frédéric Precioso received his Ph.D. in Signal and Image Processing, from University Nice-Sophia Antipolis, France, in 2004, for his work on parametric active contour segmentation methods. After one year post-doctorate as Marie-Curie Fellow at CERTH-Informatics and Telematics Institute, (Thessaloniki, Greece), where he worked on semantic methods for multimedia content understanding, he has become associate professor at the Engineering school ENSEA, Cergy, in 2005. He is interested, for ten years now, in video and image segmentation using active contours more recently focusing on bio-medical applications. Since 2005, his main topics of interest are image and video object detection and classification, interactive learning in this context, content-based image and video retrieval systems and scalability of such systems. He has been elected as associated member of the Technical Committee of IEEE Bio-Inspired Signal Processing. He is now full Professor since 2011 at the Engineering school Polytech'Nice, from University Nice-Sophia Antipolis. He is the head of KEIA group (Knowledge Extraction and Integration Algorithms) in the Laboratory I3S, UMR 7271 UNS CNRS.